

Classification of Ventricular Arrhythmia using Bat Algorithm

S. Sumathi

Department of Electrical & Electronic, Mahendra Engineering College, Namakkal, Tamil Nadu, India.

Abstract This paper presents an innovative approach to the feature extraction for reliable heart rhythm recognition. This classification of the system is comprised of three stages including data acquisition, preprocessing, feature extraction and classification of ECG signals. The first step is Preprocessing of ECGs which removes noise by using adaptive filters. Two different feature extraction methods are applied simultaneously to obtain the feature vector of ECG data. The wavelet transform is used to extract the important coefficients of the transform as the features of each ECG segment. Simultaneously, Enhanced Bat Algorithm is also applied to obtain the temporal structures of ECG waveforms. Then the Adaptive Neuro Fuzzy Inference (ANFIS) is used to classify different ECG heart rhythm. From computer simulations, the overall accuracy of classification for recognition of ventricular arrhythmia heart rhythm types reaches 99.90%.

Key Words — ECG, Denoise, Feature Extraction, Bat Algorithm, ANFIS

© 2016 mahendrapublications.com, All rights reserved

INTRODUCTION

In this competitive and flourishing technological advancement occurring across the globe, there is an essentiality for the evolution of advanced concepts in the area of Biomedical Engineering. Being the forerunner among the emerging areas in science and technology, this field demands an extensive research. The heart functioning and the diagnosis are the factors governing the ECG. This factor encircles the concepts of various multidisciplines and interdisciplines. Electrical pulses and signals that are governing components of ECG play a vital role in diagnosis and further processing. Enormous electrical equipments and electrical expertise are essential in understanding ECG processing, its execution and various allied mechanisms. The heart and the corresponding ECG are found to draw a drastic focus from electrical discipline point of view. This research is made to partially extend a valuable contribution to the field of Biomedical Engineering.

CardioVascular Disease (CVD), including heart disease and stroke, accounts for around 17 million deaths each year across the globe. According to World Health Organization (WHO) estimates, in 2003, 16.7 million people around the globe die of CVD each year. This is over 29 percent of all deaths globally. By 2020 heart disease and stroke will become the leading cause of both death and disability worldwide, with the number of fatalities projected to increase to more than 20 million a year and to more than 24 million a year by 2030 (Atlas of Heart Disease and Stroke, WHO 2004). The largest percentage increase will occur in the Eastern Mediterranean Region. The largest increase in number of deaths will occur in the South-East Asia Region. According to the most recent figures published by the Australian Bureau of Statistics and the Department of Health, cardiovascular or heart disease remains the leading cause of death in Australia, accounting for around 38% of all deaths in 2002 (Cardiovascular Health 2002, Australian Institute of Health and Welfare 2004) and 34% of all deaths in 2006. It is the leading cause of death in the USA with almost 2,000 Americans dying each day i.e. 1 death every 43

seconds.

The electrocardiogram (ECG) is regularly used in clinical practices, which describe the electrical activity of the heart. In physical checkups at the hospitals, physicians record the ECG signal after the patient has exercised to confirm his/her cardiac condition. The Holter ECG device is used for recording the ECG. The Physicians apply the device to a patient they are need to monitor his/her ECG to find the few irregular cycles in the ECG throughout the day. Physicians then read the shapes of waves and complexes. They calculate to determine the parameter whether the ECG show s cardiac disease or not. The parameters are the RR interval, PP interval, QTinterval, and ST segment. Recognition of the important points and calculations of the parameters is a tedious routine for the physician. Therefore, there is an most urgent need for an automatic ECG recognition system to reduce the burden of interpreting. The analysis of ECG signal is of much more importance in the detection of cardiac abnormalities. The most important signal intervals such as QRS complex, which is associated with the activation of electrical ventricular. The morphological classification of QRS complexes constitutes an important step in the detection of ventricular arrhythmias.

All ECG data were obtained from MIT-BIH arrhythmia database that contains records of many patients with heart troubles or abnormalities. The frequency of the ECG data was 360HZ. Each record has its respective annotation file that indicates the class of the heartbeat. A single channel ECG is collected and used to algorithm evaluation. Since there are few categories of abnormal QRS complexes in one record, we select different abnormal QRS complexes from several records. Six types of QRS complexes appeared frequently in the database. Therefore, we mainly deal with six types heartbeats which include normal beat (NORMAL), left bundle branch block beat(LBBB), right bundle branch block beat(RBBB), paced beat(PACE), premature ventricular contraction(PVC) and atrial premature contraction(APC). In the data preprocessing process, continuous ECG signals must

*Corresponding Author: sumathis@mahendra.info

Received: 20.06.2016

Accepted: 10.07.2016

Published on: 22.07.2016

S. Sumathi

be separated into many segments which contain one heartbeat. The extracted data of ECG complexes is centered around R peak. Considered that some PVC duration is great and sometimes R peak detection may be not the center of the complex, we have selected segment of 250ms before the fiducial point and 400ms after that with the R peak point is the 90th point. The R peak is detected using the Pan and Tompkins algorithm. Thus, each segment must contain one ECG heartbeat. Fig. 1 shows typical waveforms of six types of ECG segments.

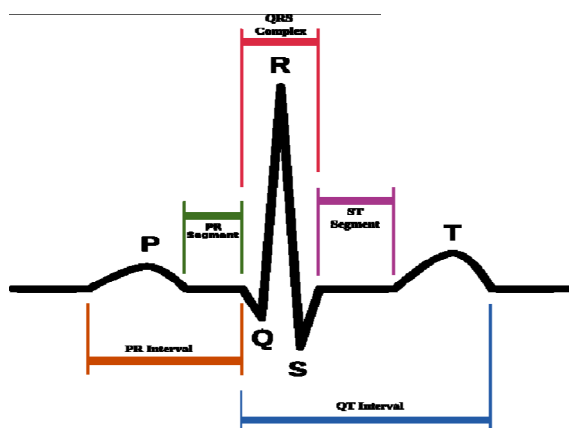


Figure 1. Typical Normal ECG Waveform

The interpretation of the ECG signal is an application of pattern recognition. The purpose of pattern recognition is to automatically categorize a system into one of a number of different classes has been pointed out by De Chazal (1998). An experienced cardiologist can easily diagnose various heart diseases just by looking at the ECG waveforms printout. The use of computerized analysis of easily obtainable ECG waveforms can considerably reduce the doctor's workload. Some analyzers assist doctor by producing a diagnosis; others provide a limited number of parameters by which the doctor can make his diagnosis by Granit (2003). After the analysis, the ECG is either interpreted as "Normal" or "Abnormal". The recognition and analyzing of ECG signal is difficult, since their size and form may change eventually and there can be considerable amount of noise in the signal. Hence, the processing of ECG is very important step in the examination of ECG by physicians.

A number of techniques have been devised by the researchers to detect QRS complex. The QRS complexes originally developed by Pan and Tompkins [1] in assembly language for implementation on a Z80 microprocessor, but it takes less time and are easily implemented. The main drawback of this algorithm is that frequency variation in QRS complexes adversely affects their performance. Therefore a real time QRS detection algorithm in the C language was developed by Hamilton and Tompkins [2]. The researchers Cuiwei Li et al [3] use the wavelet transform. It is easy to characterize the ECG waves and the QRS complex. Many automated ECG classifier systems employing different techniques for feature extraction from ECG signals and for classification of ECG patterns have been proposed by several researchers. Different solutions to ECG classification problem based on support vector machine [4], linear

discriminant analysis [5], artificial neural network [6], filter banks [7], self organizing maps [8], and wavelet transform [9], have been presented in the literature. Fuzzy set theory plays an important role in dealing with uncertainty when making decisions in medical applications. Therefore, fuzzy sets have attracted the attention and interest in modern information technology [10-13]. In this study, a new approach based on ANFIS is presented for classification of the ECG signals.

In this paper, a new approach based on Adaptive Neuro Fuzzy Inference Systems (ANFIS) is being attempted and presented for classification of the various cardiac disorders and data analysis of the ECG and by using Bat algorithm (WT) techniques. This approach is to be a very effective for the pattern recognition.

II.FEATURE EXTRACTION USING BAT ALGORITHM

The recognition of heart rhythms requires generation of the feature vector which represents the original ECG segment. A good recognition system should depend on the features representing the ECG signals in such a way, that the differences among the ECG waveforms are suppressed for the waveforms of the same type but are emphasized for the waveforms of belonging to different types of heartbeats. We perform the recognition process of heart rhythms on the single heartbeat of the ECG, proposing the description or representation by Bat algorithm model.

Bat algorithm has been developed by Xin-She Yang in 2010 [14]. The algorithm exploits the so called echolocation of bats. Bats use sonar echoes to detect and avoid obstacles. It is generally known, that sound pulses are transformed to frequency which reflects from obstacle. Bats can use time delay from emission to reflection and use it for navigation. They typically emit short loud, sound impulses. The pulse rate is usually defined as 10 to 20 times per second. After hitting and reflecting, bats transform their own pulse to useful information to gauge how far away the prey is. Bats are using wavelengths, that vary from range [0.7,17] mm or inbound frequencies [20,500] kHz. By implementation, pulse frequency and rate has to be defined. Pulse rate can be simply determined range 0 to 1, where 0 means there is no emission and by 1, bats are emitting maximum [15]. This behaviour can be used to formulate the new bat algorithm. The generalized rules for bat algorithms of all bats use echolocation to sense distance, and they also guess the difference between food/prey and background barriers in a some magical way.

III.CLASSIFICATION USING ADAPTIVE NEURO FUZZY MODEL

Adaptive Neuro Fuzzy Inference Systems (ANFIS) is one of hybrid Neuro-Fuzzy inference expert systems. Both Neural Network (NN) and Fuzzy Logic (FL) are universal estimators. They can approximate any function to any prescribed accuracy, provided that sufficient hidden neurons and fuzzy rules are available. The ANFIS is a FIS implemented in the framework of an adaptive fuzzy neural network shown in Figure 2. It combines the explicit knowledge representation of a Fuzzy Inference System (FIS) with the learning power of Artificial Neural Networks (ANN) [16-18]. The advantage of a fuzzy set is the representation of prior knowledge into a set of constraints to reduce the

optimization research space is utilized. The adaptation of back propagation to structured network to automate fuzzy control parametric tuning is utilized from neural network. For precise parameters that define membership functions, ANFIS employs gradient descent algorithm to fine-tune them. For subsequent parameters that define the coefficients of each equation, ANFIS uses the least-squares method to identify them. This approach is thus called hybrid learning method since it combines gradient descent algorithm and least-squares method. Functionally there are no constraints on the node functions of an adaptive network except for the requirement of piecewise differentiability. Hybrid systems are a growing research area in medical applications. This system overcomes some of the major drawbacks of conventional expert systems such as the consultation with human experts for knowledge acquisition. A Neuro Fuzzy contains three main components, which are fuzzification stage, the rule base and the defuzzification stage. The first layer represents the direct input layer. Each node represents one ECG feature and sends these input variables from the features extraction to next layer directly. The second layer is the fuzzification layer. In this layer, each extracted ECG feature is fed to fuzzy nodes. The third layer represents the fuzzy rules stage. Each node in this layer represents one fuzzy represents one fuzzy if-then rule. The last layer is the neural network connecting layer.

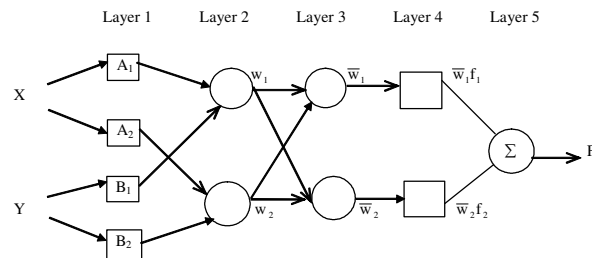


Figure 2. ANFIS architecture for a two rule sugeno system

IV. ECG DATABASE

Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia Database is used for the present work. There are forty-eight groups of patients in MIT-BIH Database. Each group data include two ID files, one sample data file and one annotation file. The annotation file records the name, ECG data and length in each group dataset. The lead II and lead V1 were stored in the sampling data file which was sampled at 360Hz with 11 bit resolution. The ECG recording time is about thirty minutes and data length about 2.5Mbytes [19].

V. RESULTS & DISCUSSION

The block diagram of the proposed method for ECG beat classification is depicted in Figure 3. The method is divided into three steps: (1) pre-processing (2) ECG Preprocessing and (3) Feature Extraction using BAT Algorithm (4) classification by ANFIS.

Step 1: Pre-processing

Clinical ECG recordings undergo several stages of filtering in an attempt to reduce noise. Each noise source resides in a characteristic frequency band. Poor conductance between skin and electrode creates slowly varying potentials which

manifests as baseline wander in the ECG. The output function of this wavelet transform will be our filtered signal.

$\Psi(t)$ given in the equation (1). The parameters of this filtering are the attenuation factor β , and the basic frequency f .

$$\Psi(t) = \exp\left(-\frac{t^2}{\beta^2}\right) \cos(2\pi ft) - \lambda \quad (1)$$

β - Attenuation factor, f - base frequency, λ - DC factor eliminator, $\Psi(t)$ -Wavelet Transform

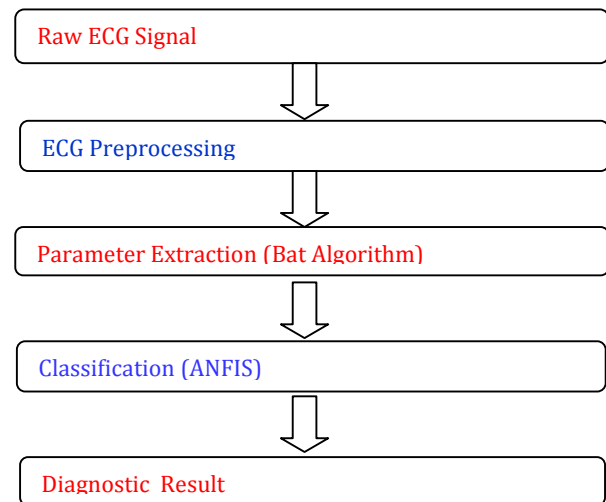


Figure 3. Proposed method for ECG beat classification

Our goal is to find those parameter values, which contribute the most to a good QRS detection ratio. The value for parameters will be chosen so that it assures all the significant parts of the sum defined in Equation (1), but it eliminates all additional calculations for the attenuated part of signal, that hardly influences the result. Elimination of low and high frequency components from the signal using the filters is the concrete effect of the wavelet filtering shown in Figure 4, Figure 5.

Step 2 : Denoising of ECG data is a preprocessing step that removes noise and makes ECG file useful for subsequent steps in the algorithm. The Sgolay FIR smoothing filter is used for filtering.

The R peaks of the ECG signal are detected as shown in Fig.6. Distance between two R peaks is called RR interval. 2/3 samples to the right and 1/3 samples to the left of R peak is considered as one beat duration as given Fig7.

Step 3: ECG feature selection using BAT algorithm

The feature extraction methods generally yield a large number of features, and many of these might be insignificant. Therefore, the effective technique in this study is to extract the key features useful in the classification of ECG beats. This present work, 25 key features from 400 samples is extracted using Bat algorithms shown in Figure 5,6,7.

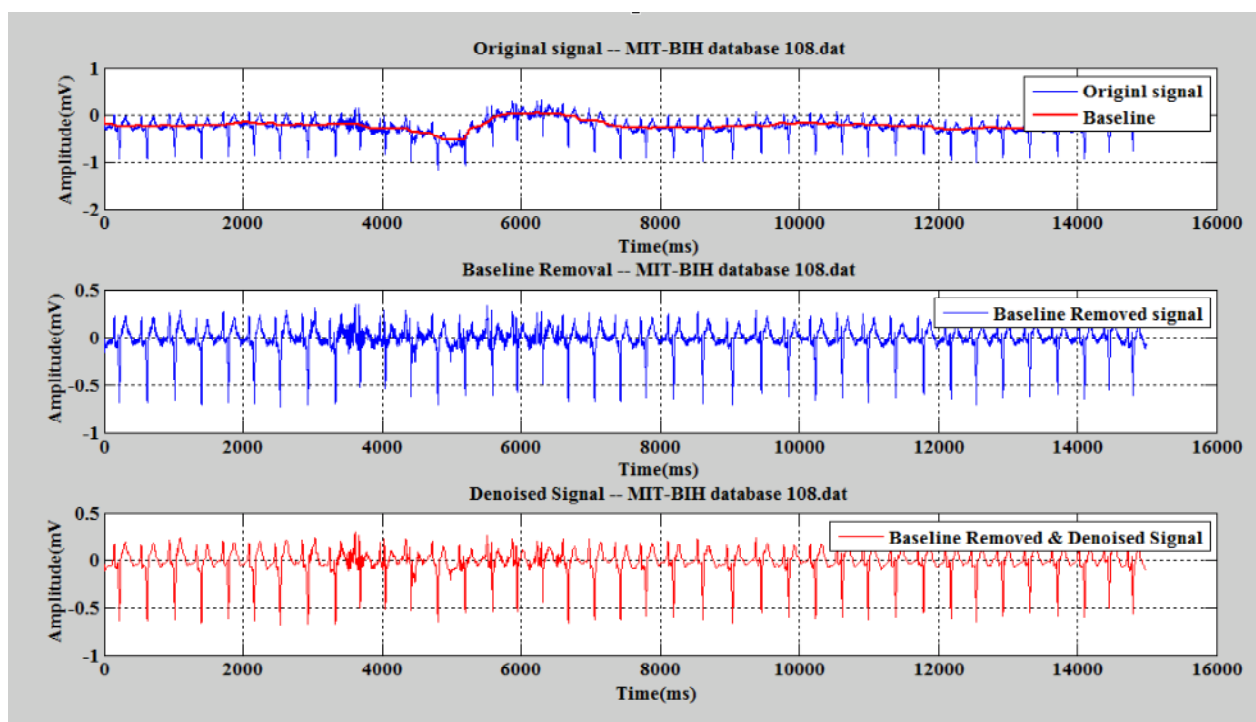


Figure 4. Baseline removed and Denoised signal

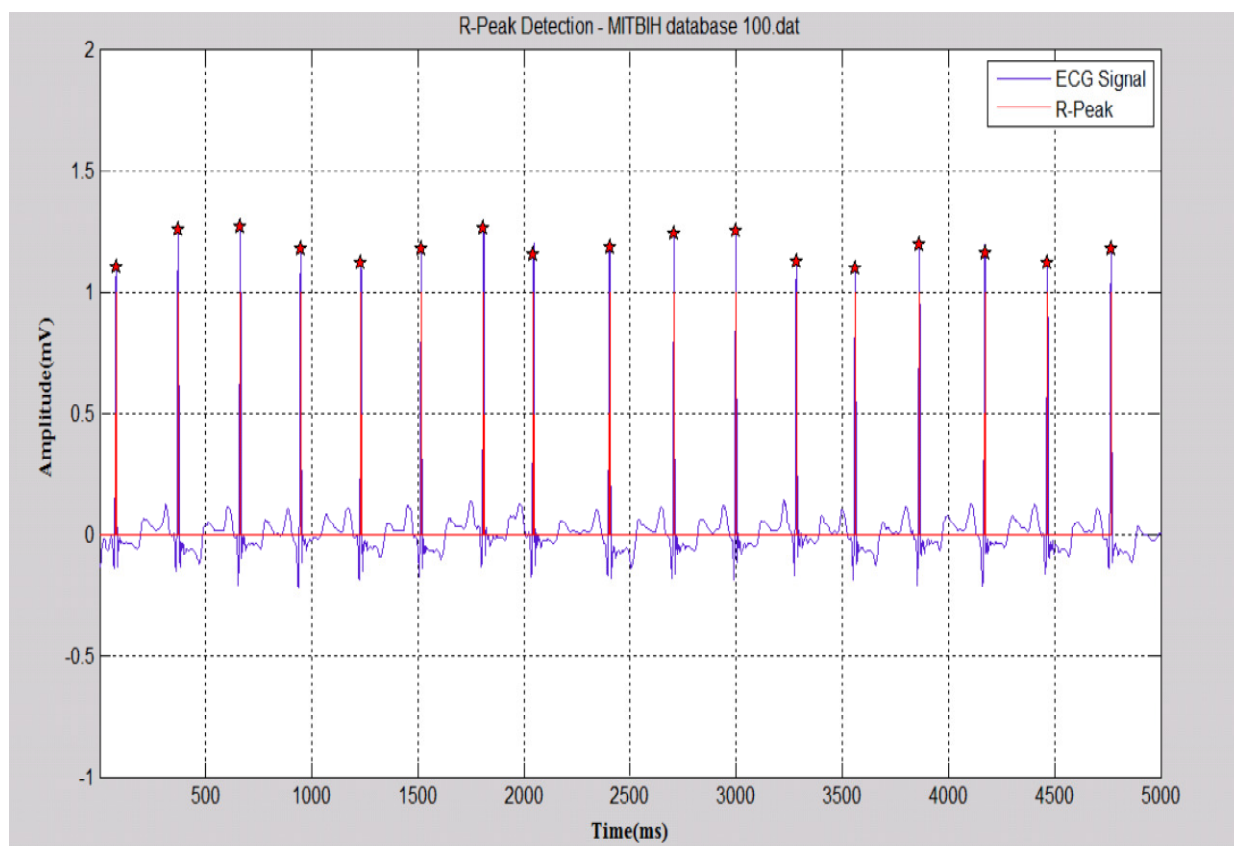


Figure 5. R peak detection

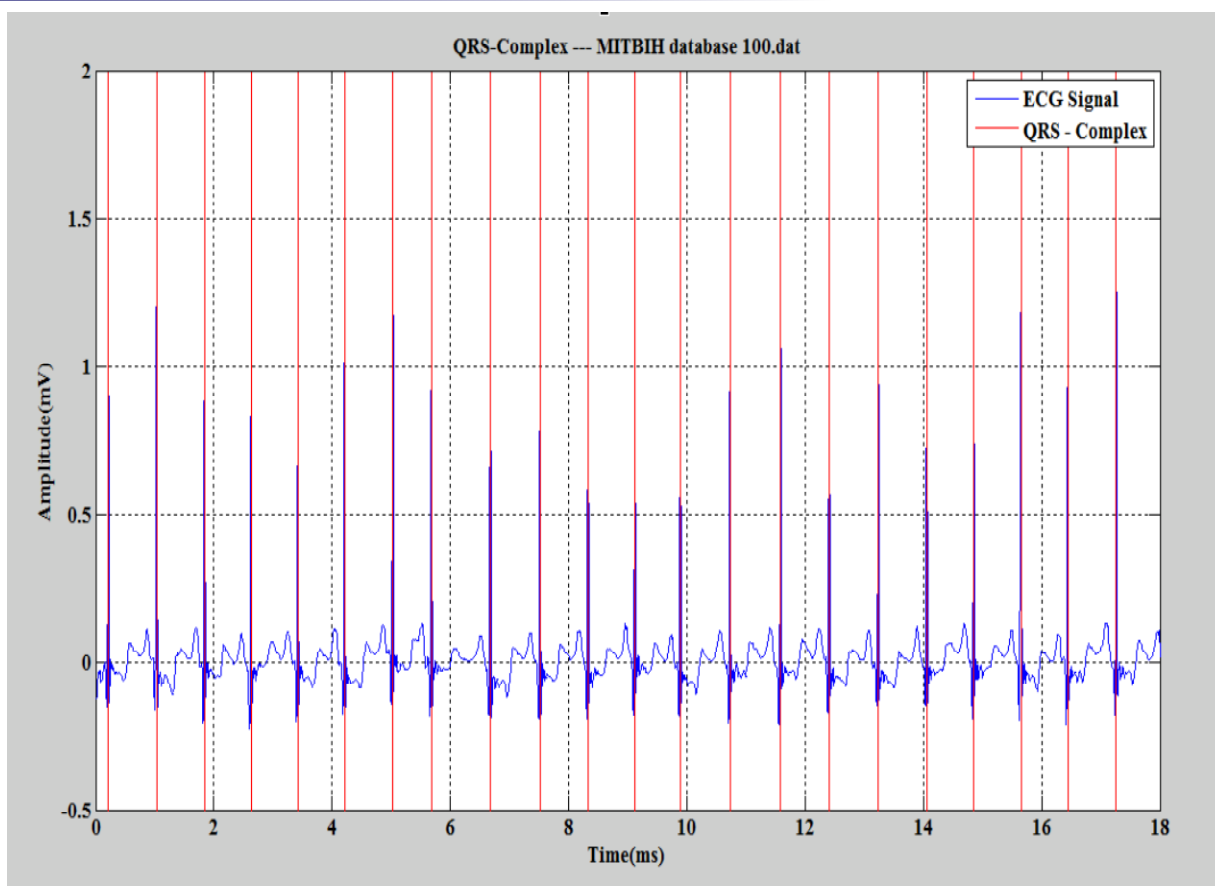


Figure 6 QRS Complex

The scale towards the right comprises of various colours depicting the range of energy available in the particular sector of waveform after the elimination of baseline. The Sensitivity reports the percentage of the beats that were correctly detected by the algorithm.

$$\text{Sensitivity (\%)} = \frac{TP}{TP + FN} \quad (2)$$

where TP is True Positive (Correctly detected)

FN is False Negative (No. of missed detection)

The Positive Predictivity reports the percentage of beat detections which were true beats.

$$\text{Positive Predictivity(\%)} = \frac{TP}{TP + FP} \quad (3)$$

FP is False Positive (No. of extra detection) in QRS complex Table I shows the results of the Symlet Spline Wavelet based QRS detection algorithm for the MIT-BIH Database. The average sensitivity of the algorithm is 99.56% and its Positive Predictivity is 99.90%.

TABLE I. Results of the bat algorithm based QRS detection algorithm for the MIT-BIH Database

Record	TP	FN	FP	Sensitivity (%)	Positive Predictivity (%)
100	2261	11	0	99.51	100
101	1856	7	4	99.62	99.78
102	2185	1	0	99.95	100
103	2079	3	0	99.85	100
104	2215	15	43	99.32	98.92
105	2543	11	4	99.56	99.84
109	2526	4	3	99.84	99.88
111	2126	1	0	99.95	100.00
113	1794	0	1	100.00	99.94
114	1878	1	3	99.95	99.84
115	1941	5	1	99.74	99.94
116	2403	14	6	99.42	99.75
117	1538	0	1	100.00	99.93
Total	52282	237	277	99.56	99.90

The FN and FP values tend to dropdown abruptly to lower value for records corresponding to (104, 203 and 207). The ECG waveforms in the above records are characterized by high complexity which leads to internal difficulties in detecting the QRS complex.

Step 4: Classification of Life-Threatening Arrhythmia using ANFIS

The life threatening arrhythmia by the application of the strategic problem solving capability ANFIS. The classification of cardiac arrhythmia using ANFIS uses GBELLMF. GBELLMF (x, parameters) returns a matrix, which is the generalized bell membership function evaluated at x. Parameters are a 3-element vector that determines the shape and position of this Membership Function (MF). The Generalized BELL membership function is shown in Fig. 7. It depends on three parameters a, b, and c as given by the Equation (4)

$$f(x, a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (4)$$

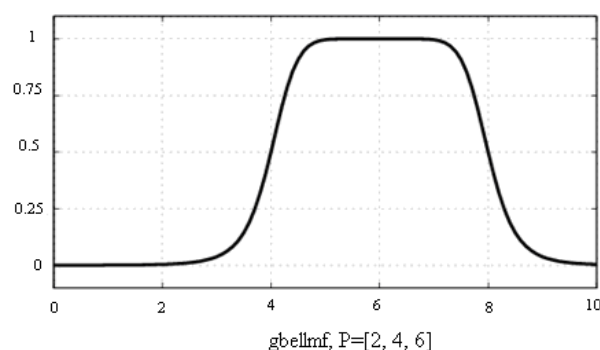


Figure. 7. Generalized BELL membership function

The whole process is to identify the shapes of QRS complex. The Relationship between SUB-ANFIS and the shape of ECG segments showed in Table II

TABLE II: Relationship between Sub-ANFIS And the Shape of ECG Segments

SUB-ANFIS	ANFIS 1	ANFIS 2	ANFIS 3	ANFIS 4	ANFIS 5	ANFIS 6
The shape of the ECG Segment	SVT	PVC	VF	VT	VFLU	MI

Data were collected of the shapes of P wave and QRS complex and divided into two subsets. One is called as training data and the other as test data. The data is further subdivided as output=1 and output=0. This forms the

training data. The classification of the NSR, PVC, VT, VF, VFL, AF arrhythmia is carried out using a ANFIS whose input is the energy level calculated by the wavelet transform as described above. To achieve clustering of the ECG waveforms into five different classes, the number of output layer should be five as shown in Table III. A target vector was arranged as the desired output for each class. It is a set of Boolean value vectors. Accompanying each record in the MIT-BIH Database there is an annotation file in which each heartbeat has been identified by expert cardiologist annotators. This annotated information can be used for designing the target vector and evaluating the classifier performance.

TABLE III :FIVE OUTPUT TARGET VECTOR

Signal	Vector	Heart Condition
SVT	[1 0 0 0 0 0]	supraventricular tachycardia
PVC	[0 1 0 0 0]	Preventricular Contraction
VT	[0 0 1 0 0]	Ventricular Tachycardia
VF	[0 0 0 1 0]	Ventricular Fibrillation
VFL	[0 0 0 0 1 0]	Ventricular Flutter
AF	[0 0 0 0 0 1]	Myocardial Ischemia

The data used for training and testing in this research is obtained from the MIT-BIH Database and was extracted by the feature extraction algorithm. ECG signal classifier becomes, more complex as the number of signal abnormalities increase. In this study, 70% of data set was to be training and 30% of data set was to be testing data. The highest accuracy was obtained by dividing the data for training and testing. The training data set was used to train the ANFIS model, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained ANFIS model for classification of the four classes of ECG signals. Each ANFIS used 430 training data in 840 training epochs and the step size for parameter adaptation had an initial value of 7.5038×10^{-5} . After training, 185 testing data were used to validate the accuracy of the ANFIS model for classification of the ECG signals.

Higher accuracy was obtained by dividing the data into two equal parts for training and testing. The test performance of the classifiers can be determined by the computation of total classification accuracy of each type of beat using the following Equation (5)

(5)The performance results of the ANFIS used for classification of the ECG signals is given in Table IV.

TABLE IV. PERFORMANCE RESULTS OF ANFIS

S. Sumathi

Type of Arrhythmia	Training Data	Testing Data	Misclassification	Classification Accuracy (%)
Supraventricular tachycardia	95	45	0	100.00
Preventricular Contraction	80	35	1	97.14
Ventricular Tachycardia	75	30	1	94.33
Ventricular Fibrillation	95	40	1	97.00
Ventricular Flutter	85	35	1	94.11
Myocardial Ischemia	80	34	2	97.52
Total	430	185	7	98.24

The presented ANFIS model combined the neural network adaptive capabilities and the fuzzy logic quantitative approach. Some performance concerning the salient features on classification of the ECG signals were obtained through analysis of the ANFIS. The classification results and statistical measures were used for evaluating the ANFIS. The total classification accuracy of the ANFIS model was 96.34% as presented in Table V.

VI.CONCLUSION

The experiments of recognition ventricular of arrhythmias and normal beat were carried out on MIT-BIH Arrhythmia Database. The ANFIS used for classification of the ECG beat was trained, cross validated and tested with the extracted features from Bat Algorithm and ANFIS of the ECG signals. Computer simulations showed that our approach gives the excellent performances of successful recognition, accuracy is found to be 99.90

REFERENCE

- [1] Pan,J., Tompkins, W.J., 1985.A real-time QRS detection algorithm. IEEE Trans. Biomed Eng.,32, pp. 230–236.
- [2] Hamilton,P.S., Tompkins,W.J.,1986 .Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans. Biomed. Eng., 33, 1157–1165.
- [3] Cuiwei. Li., C. Zheng, and. Tai, C., 1995.Detection of ECG characteristic points using wavelet transform. IEEE Trans Biomed Eng, 42(1), 8-21.
- [4] Qibin Zhao., Liqing Zhang,2005. ECG feature extraction and classification using wavelet transform and support vector machines. IEEE Trans. Biomed. Eng., 23, 1089-1092.
- [5] Lee,L.,Park,K.L.,Song,M.H., Lee, K.J., 2005.Arrhythmia classification with reduced features by linear discrimination analysis. IEEE EMBS,27th Annual International Conference, 1142-1144.
- [6] Louis, C Pretorius,Cobus Nel., 2002.Feature Extraction From ECG for Classification By Artificial Neural Network. University of Pretoria, In IEEE Proceeding, 67-70.
- [7]Afonso,X.,Tompkins,W.J., Nguyen.,Luo., 1999.ECG beat detection using filter banks. IEEE Transactions on Biomedical Engineering, 46 (2) 192-202,
- [8]Talbi., Charef, A., 2009.PVC discrimination using the QRS power spectrum and self –organizing maps. Science Direct, Computer Methods and programs in medicine, pp 45-52.
- [9] Mahmoodabadi, S. Z., Ahmadian, A.,Abolhasani, M. D., Eslami., Bidgoli,J. H., 2005.ECG Feature Extraction Based on Multiresolution Wavelet Transform. IEEE EMBS, 27th Annual International Conference, 1142-1144.

- [10] Anuradha, B., Suresh Kumar,K,Veera Reddy,V. C., 2008.Cardiac Arrhythmia classification Fuzzy Clasifiers. Journal of Theoretical and applied Information Technology, 4,353-358.
- [11] Benali,R., Chikh, M.A.,2009.Neuro-fuzzy classifier classifier for cardiac arrhythmias recognition.Journal of Theoretical and applied Information Technology, 4, 577-582.
- [12]Kannathal,N.,Sadasivan,K.,Puthusserypady.,2005.Cardiac State Diagnosis using Adaptive Neuro-Fuzzy Technique.Proceedings of theIEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, 38643867.
- [13] Emran, M.Tmil., Hafeezah Kamarudin., 2009.Heartbeat Electrocardiogram signal feature extraction using discrete wavelet transform.,in Proceedings of CSPA.1112-1117.
- [14]Yang,X.S.,2010.A new metaheuristic bat-inspired algorithm. Nature Inspired Cooperative Strategies for Optimization (NISCO 2010), 65–74.
- [15] Sumathi,S.,Sanavullah,M.Y., 2009.Comparative Study of QRS Complex Detection in ECG Based on Discrete Wavelet Transform. International Journal of Recent Trends in Engineering, 2(5), 273-277.